Autonomous Robotics

Robots and LLMs

Josiah Hanna
University of Wisconsin — Madison

Announcements

Final project released today.

Learning Outcomes

After today's lecture, you will:

- Understand what it means for a robot's knowledge to be grounded.
- Understand how LLMs can be used in robotics and how they go beyond other learning approaches.

Motivation for LLMs in Robotics

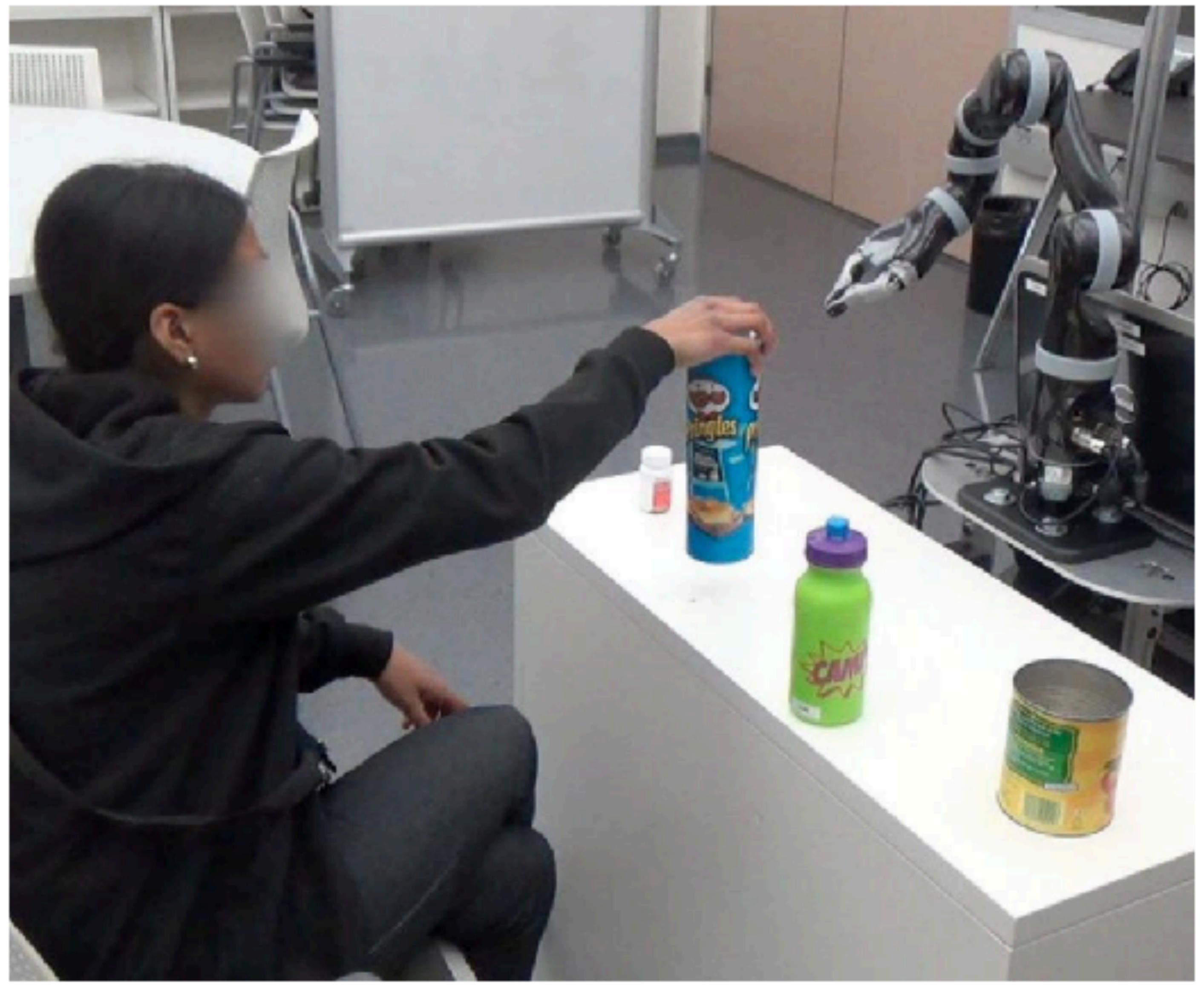
- Provide robots with common-sense reasoning.
- Allow robots to learn from repositories of human knowledge (i.e., the internet)

Grounding

- Knowledge is mapped to the physical senses and actions of a robot.
  - Example: “Pick up the red cup.”

What sequence of controls will accomplish this instruction?

What does a red cup look like in a camera image?



Learning Multi-Modal Grounded Linguistic Semantics by Playing “I Spy”

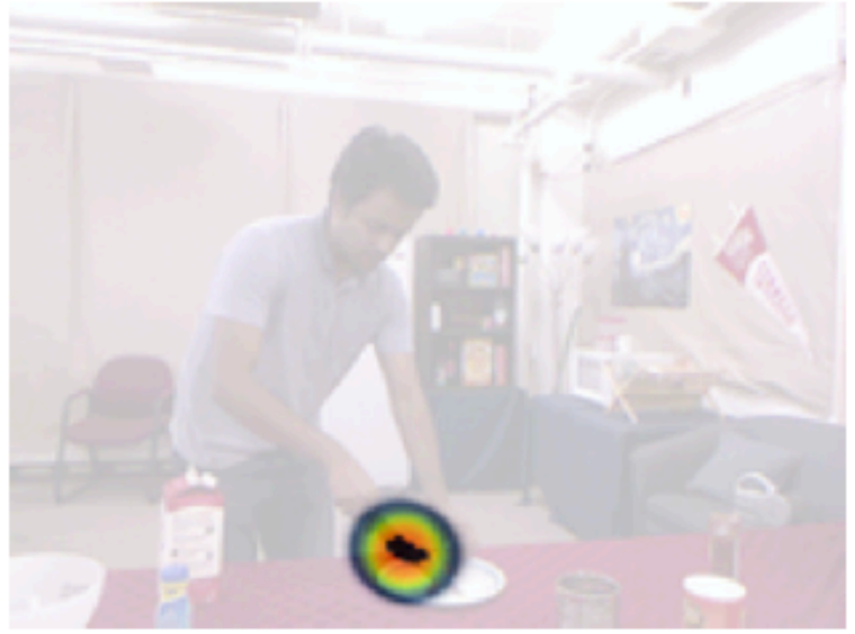
Prior Work (2014): RoboBrain

'Robo Brain' mines the Internet to teach robots

By Bill Steele

August 25, 2014

The **standing_human** can **cut** using a **knife** as shown in **$heatmap_2.**



Anticipation http://pr.cs.cornell.edu/anticipation/

236 3

[Robo Brain: Large-Scale Knowledge Engine for Robots]

Large Language Models

- Large neural networks that take text input and generate text output.
  - Formally, model $p(\texttt{nextword} \mid \texttt{prompt} + \texttt{previouswords})$.
  - Generate outputs one word at a time by sampling from this distribution.
  - The basic training procedure is self-supervised classification: take segments of words from a large text corpora and predict the next word that follows.
- Old idea but now **significantly** scaled up with internet scale data, datacenters of GPUs, and the transformer neural network architecture.

Language Models

- Basic idea: use probabilistic models to **assign a probability to a sentence:**

$$P(W) = P(w_1, w_2, \ldots, w_n) \text{ or } P(w_{\text{next}} | w_1, w_2 \ldots)$$

- Goes back to Claude Shannon
  - "Father of Information Theory"
  - Information theory: letters

Zero-order approximation	XFOML RXKHRJFFJUJ ALPWXFWJXYJ FFJEYVJCQSGHYD QPAAMKBZAACIBZLKJQD
First-order approximation	OCRO HLO RGWR NMIELWIS EU LL NBNESEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL
Second-order approximation	ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE
Third-order approximation	IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE
First-order word approximation	REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT NATURAL HERE HE THE A IN CAME THE TO OF TO EXPERT GRAY COME TO

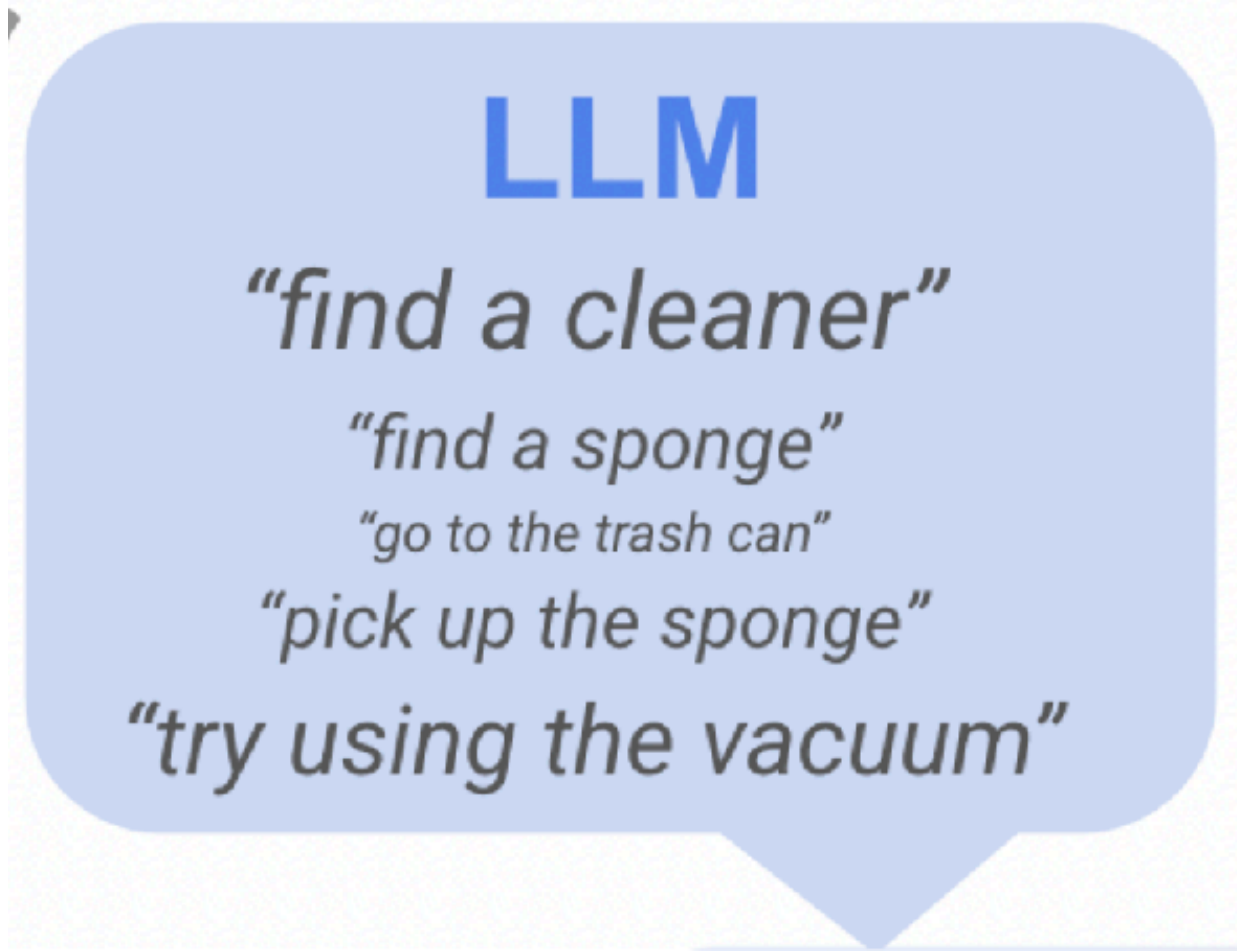
Large Language Models (cont'd)

- Self-supervised pre-training doesn't produce a highly useful model.
- So typically follow with some type of fine-tuning:
  - Supervised fine-tuning: human annotators provide better outputs and model imitates those (imitation learning).
  - RLHF: human annotators rank responses, learn a reward function (IRL), and then use RL with the learned reward.

Ungrounded LLMs

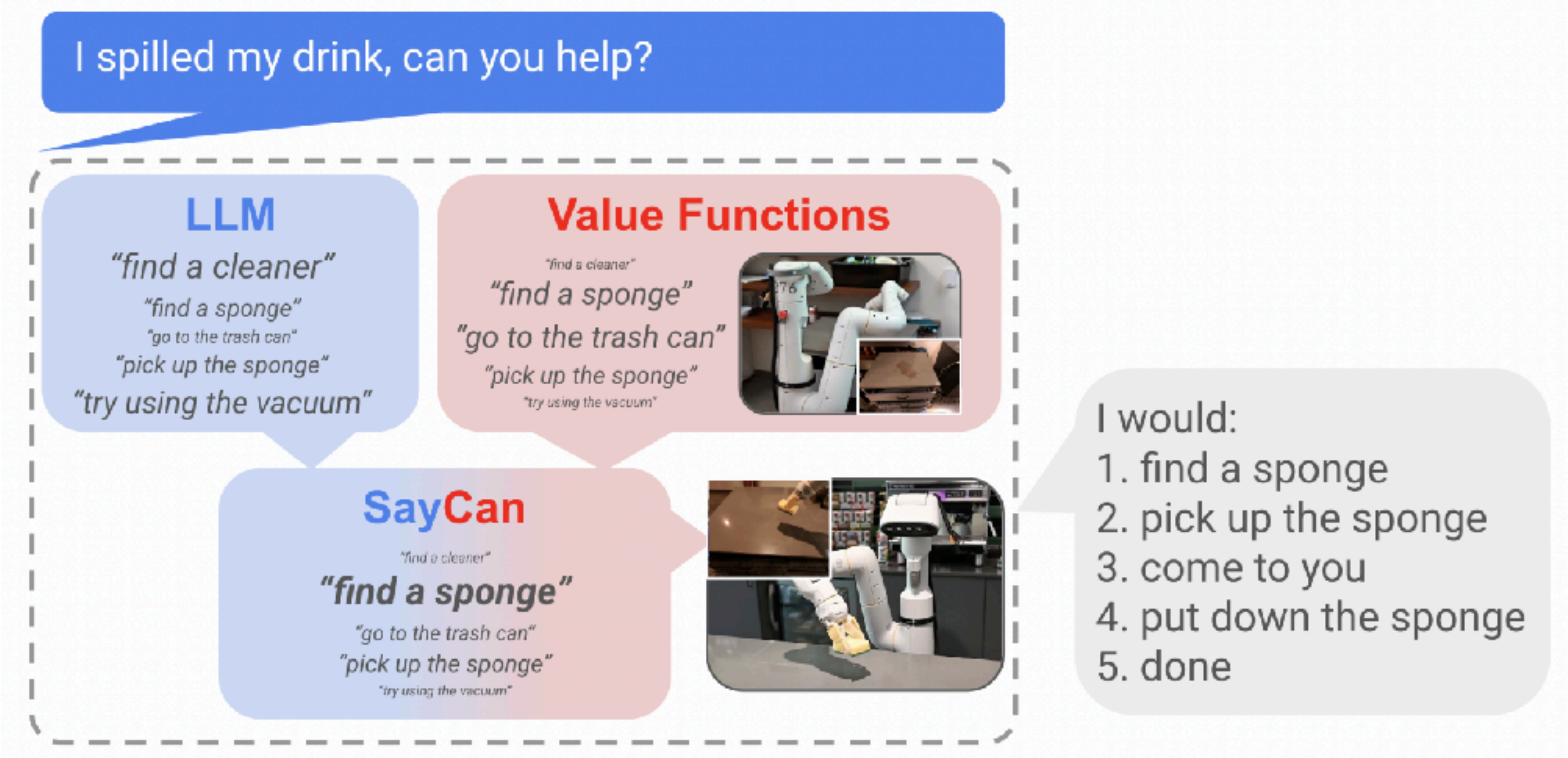
- No matter how sophisticated the outputs of an LLM are, they are not grounded in physical experience and ability.

- VLM: vision-language models.



SayCan

- LLMs can identify potential skills to complete a task.
- Value functions predict probability of a skill succeeding in a state.
- SayCan: combine both to identify skills with a high probability of task success.



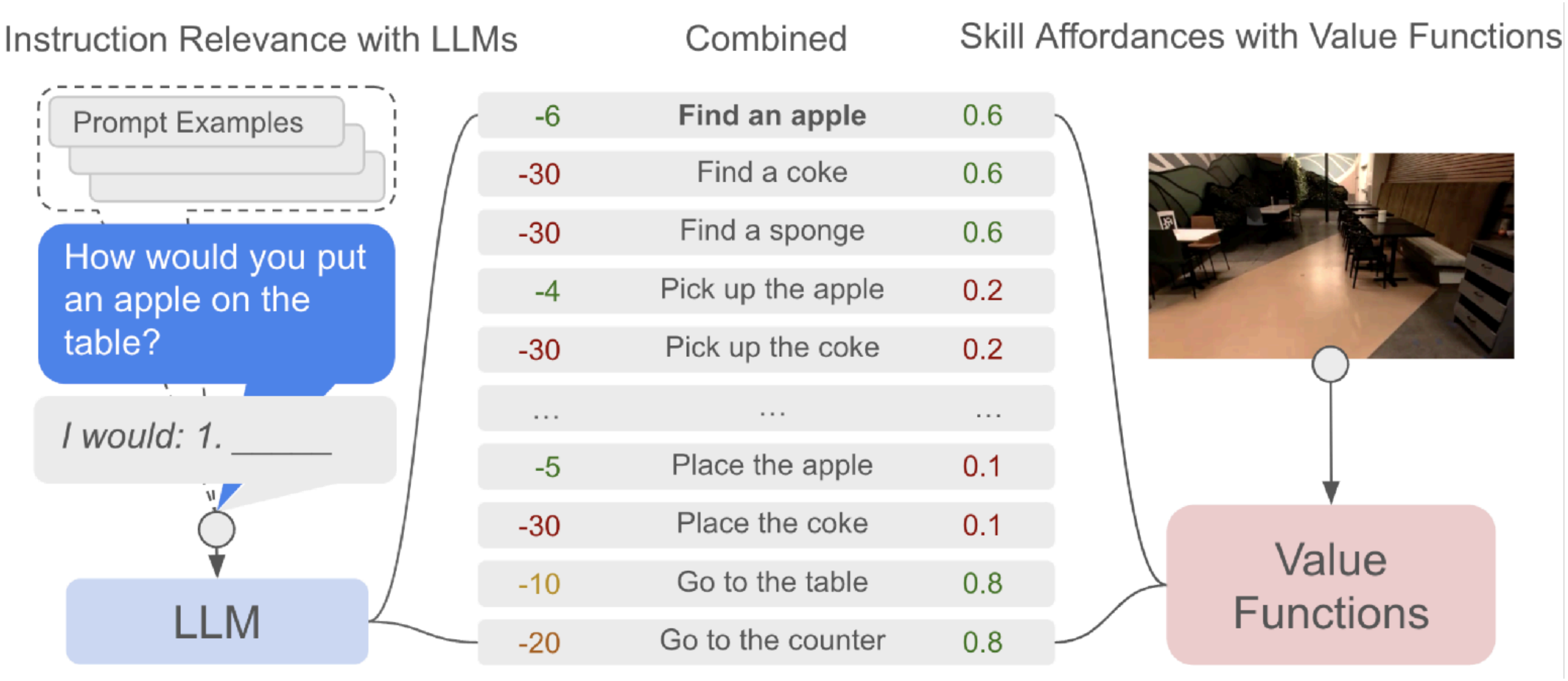
SayCan

- First, assume that we already have a skill library, Π, and for $\pi \in \Pi$ we know the probability that π can be successfully executed, $p(\texttt{success} \mid s, \ell_\pi)$.
  - "If I ask robot to do ℓ_n, will it do it?"
  - Called a value or affordance function.
- When given a prompt, i, for a new task, the robot scores all skills with $p(\texttt{success} \mid s, \ell_\pi) p(\ell_\pi \mid i)$ and takes the skill most likely to succeed.

Value Function Training

- Skills in the skill library come from either behavior cloning or reinforcement learning.
  - Each skill is a policy ($\pi : S \rightarrow A$) and also has a short text description, ℓ_n.
- Now, given a skill, we need to learn $p(\texttt{success} \mid s, \ell_n)$.
  - Equivalent to $v_\pi(s)$ for an MDP where the reward is zero except for upon success.
  - Learn the success probability with temporal difference learning.

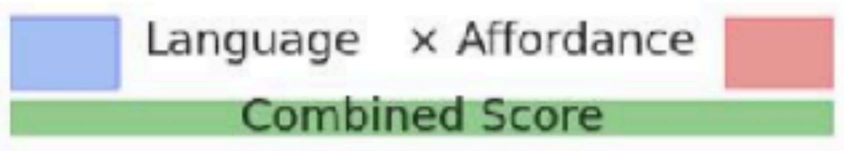
SayCan

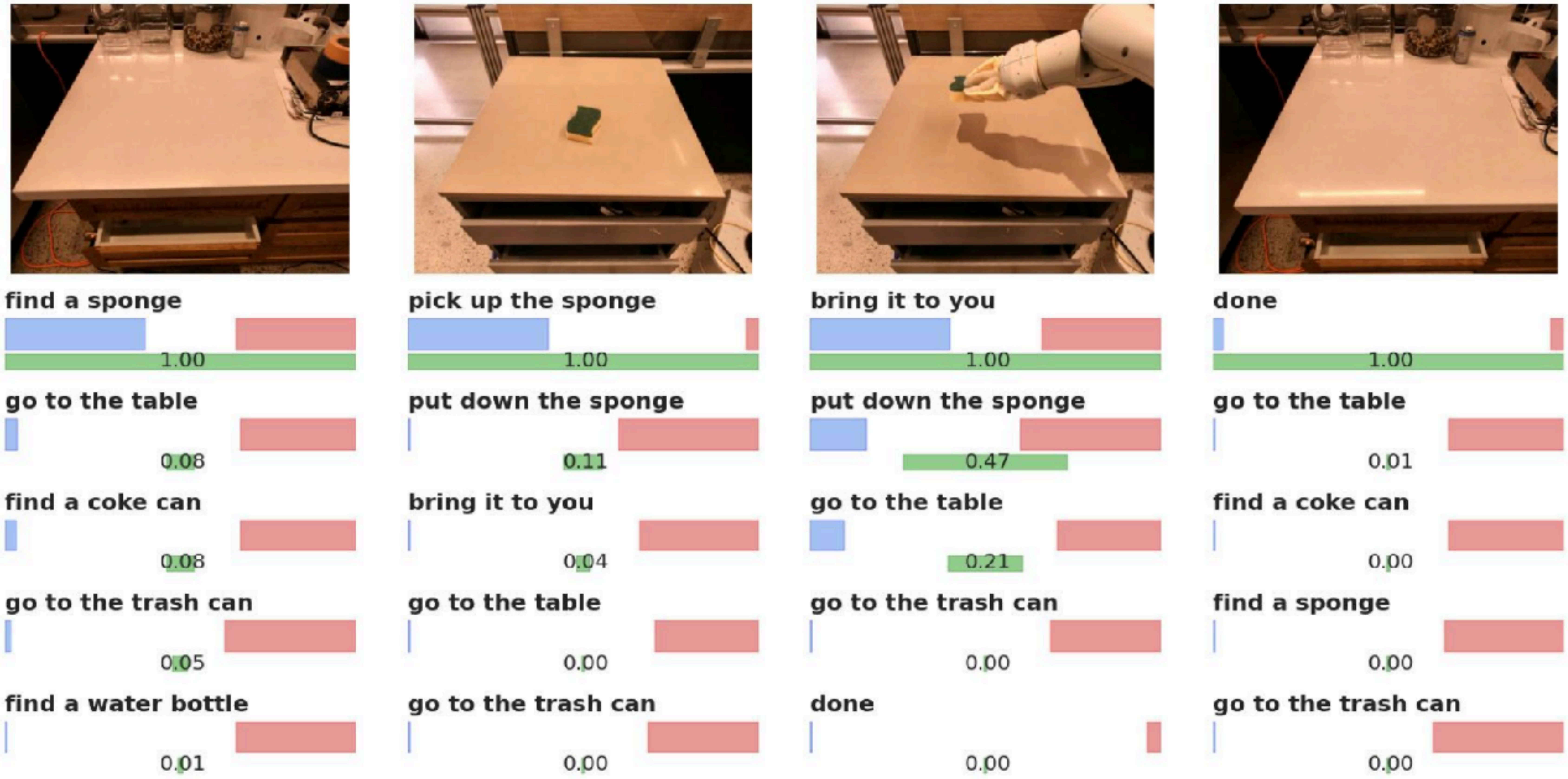

Instruction Relevance with LLMs
Combined
Skill Affordances with Value Functions
Prompt Examples
How would you put an apple on the table?
I would: 1. ______
LLM
-6 Find an apple 0.6
-30 Find a coke 0.6
-30 Find a sponge 0.6
-4 Pick up the apple 0.2
-30 Pick up the coke 0.2
... ...
-5 Place the apple 0.1
-30 Place the coke 0.1
-10 Go to the table 0.8
-20 Go to the counter 0.8
Value Functions

SayCan

Human: I spilled my coke, can you bring me something to clean it up?

Robot: I would
1. Find a sponge
2. Pick up the sponge
3. Bring it to you
4. Done

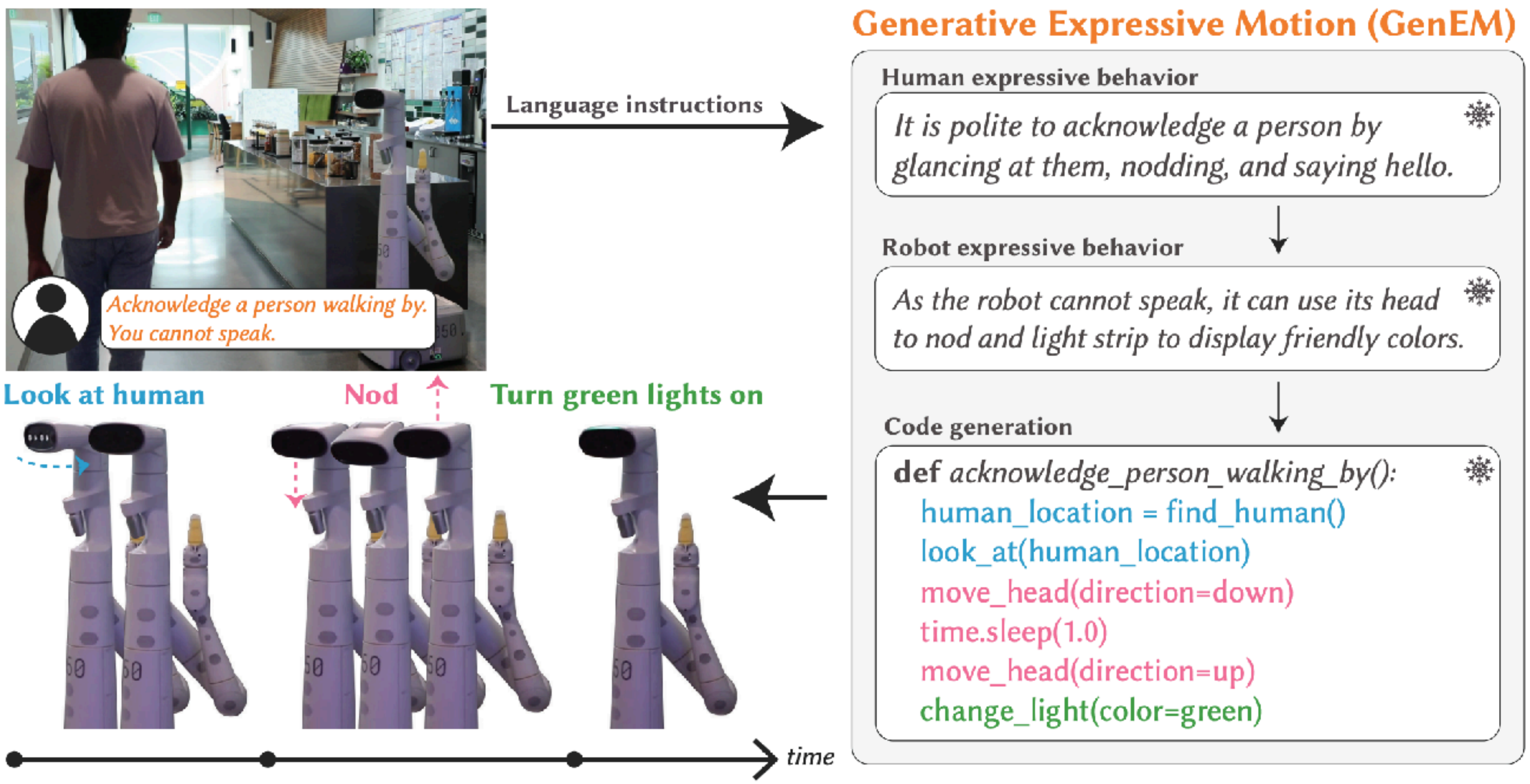




LLMs and Robots

- Strengths:
    - Take advantage of internet-scale data to help robots understand the world.
    - LLMs have demonstrated remarkable capabilities resembling advanced reasoning and problem solving —> use to inform robot action.
- Weaknesses & open questions:
    - Large models have a high storage footprint and inference cost.
    - Models may hallucinate.
    - Robots have a different embodiment than humans; knowledge may not transfer.
    - Continual learning?

Code Generation for Robots



Generative Expressive Robot Behaviors using Large Language Models

Summary

Today we covered:

1. What it means for language to be grounded.

2. Why LLMs are not grounded.

3. The SayCan approach as an example method for grounding advanced LLMs.

Action Items

Human-robot interaction readings for next week.

Begin final project.